

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 31, 2003, 13:28:27 ; Search time 104 seconds
(without alignments)
1305.151 Million cell updates/sec

Title: US-10-082-894-2

Perfect score: 2786

Sequence: 1 MDKYNQVQKVCVWIDWG.....LMGLPVPPENDGVPLLEORG 526

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:**

- 1: sp_archaea:**
- 2: sp_bacteria:**
- 3: sp_fungi:**
- 4: sp_human:**
- 5: sp_invertebrate:**
- 6: sp_mammal:**
- 7: sp_mhc:**
- 8: sp_organelle:**
- 9: sp_phase:**
- 10: sp_plant:**
- 11: sp_rodent:**
- 12: sp_virus:**
- 13: sp_vertebrate:**
- 14: sp_unclassified:**
- 15: sp_rvirus:**
- 16: sp_bacteriap:**
- 17: sp_archaeap:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | ID | Description |
|------------|--------|---------------|--------|----|-------------|
| 1 | 1900 | 68.2 | 521 | 5 | Q8IA68 |
| 2 | 1900 | 68.2 | 539 | 5 | O44742 |
| 3 | 1398 | 50.2 | 520 | 3 | Q9HG4 |
| 4 | 1075.5 | 38.6 | 526 | 16 | Q8DCW1 |
| 5 | 1059 | 38.0 | 523 | 16 | Q8FCA6 |
| 6 | 992 | 35.6 | 505 | 16 | Q8CPV4 |
| 7 | 971.5 | 34.9 | 503 | 5 | Q8SR33 |
| 8 | 966.5 | 34.7 | 506 | 16 | Q8EW33 |
| 9 | 742 | 26.6 | 559 | 10 | Q9XE59 |
| 10 | 729 | 26.2 | 559 | 10 | Q9ZS53 |
| 11 | 727 | 26.1 | 560 | 10 | Q8LF55 |
| 12 | 715 | 25.7 | 559 | 10 | Q9SDL3 |
| 13 | 714.5 | 25.6 | 560 | 10 | Q94DY7 |
| 14 | 706 | 25.3 | 557 | 10 | Q8LB62 |
| 15 | 700.5 | 25.1 | 557 | 10 | Q8HI61 |
| 16 | 700 | 25.1 | 557 | 10 | Q94AY0 |

| | | | | | |
|----|-------|------|------|----|--------|
| 17 | 699 | 25.1 | 557 | 10 | Q932F2 |
| 18 | 659 | 23.7 | 557 | 10 | Q94KV1 |
| 19 | 639.5 | 23.0 | 516 | 10 | Q942J3 |
| 20 | 626.5 | 22.5 | 550 | 5 | Q9NG18 |
| 21 | 189.5 | 6.1 | 1011 | 16 | Q8EX30 |
| 22 | 167.5 | 6.0 | 105 | 3 | Q9YH2 |
| 23 | 119 | 4.3 | 430 | 5 | Q19519 |
| 24 | 115.5 | 4.1 | 545 | 16 | Q8DC63 |
| 25 | 110 | 3.9 | 709 | 5 | Q21376 |
| 26 | 110 | 3.9 | 927 | 5 | Q9V4N0 |
| 27 | 109 | 3.9 | 685 | 16 | Q8F9T4 |
| 28 | 109 | 3.9 | 1946 | 16 | Q8YTN2 |
| 29 | 108 | 3.9 | 509 | 11 | Q88548 |
| 30 | 107 | 3.8 | 546 | 16 | Q8EBO9 |
| 31 | 106 | 3.8 | 1088 | 10 | Q8SJF0 |
| 32 | 106 | 3.8 | 1424 | 10 | Q9FF69 |
| 33 | 105.5 | 3.8 | 921 | 10 | Q9XHG0 |
| 34 | 105 | 3.8 | 874 | 5 | Q8SR58 |
| 35 | 104.5 | 3.8 | 477 | 17 | Q8PTD6 |
| 36 | 104 | 3.7 | 715 | 10 | Q8L6C6 |
| 37 | 104 | 3.7 | 790 | 13 | Q91838 |
| 38 | 104 | 3.7 | 1543 | 3 | Q8X023 |
| 39 | 103.5 | 3.7 | 390 | 16 | Q97HE6 |
| 40 | 103 | 3.7 | 588 | 16 | Q8FF97 |
| 41 | 103 | 3.7 | 773 | 2 | Q93FB8 |
| 42 | 103 | 3.7 | 4840 | 2 | Q93HJ4 |
| 43 | 102.5 | 3.7 | 547 | 10 | Q9MA11 |
| 44 | 102 | 3.7 | 494 | 16 | Q8CWB5 |
| 45 | 102 | 3.7 | 913 | 4 | O75061 |

ALIGNMENTS

RESULT 1

Q8IA68 PRELIMINARY; PRT; 521 AA.
 ID O8IA68;
 AC O8IA68;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Hypothetical protein F57B10.3b.
 GN F57B10.3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Bristol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA Waterston R.;
 RT "Genome sequence of the nematode C. elegans: a platform for
 investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Bristol N2;
 RA Greco T., Elliott G., Keppler D.;
 RT "The sequence of C. elegans cosmid F57B10.";
 RN [3]
 RP Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 RP SEQUENCE FROM N.A.
 RC STRAIN-Bristol N2;
 RA Waterston R.;
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF039713; AA012419.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 521 AA; 57110 MW; D13B64A84C99EC83 CRC64;

Query Match

Best Local Similarity 68.2%; Score 1900; DB 5; Length 521;

Matches 352; Conservative 76; Mismatches 86; Indels 14; Gaps 5;

DR ProDom: PD004429; Pgm_bpd_ind: 1.
DR TIGRFAMs: TIGR01307; pgm_bpd_ind: 1.
SQ SEQUENCE 520 AA; 57289 MW; D2BDA4A4DB3B12BB2 CRC64;

Query Match 50.2%; Score 1398; DB 3; Length 520;
Best Local Similarity 53.4%; Pred. No. 2e-106;
Matches 281; Conservative 79; Mismatches 144; Indels 22; Gaps 7;

Qy 7 VQKVLVVIDGWLGS--DEQGNATAKAKTPIM-----DKLSCSNWOKLEAHGLHVG 57
Db 4 VQKVLVVIDGWLGS--DEQGNATAKAKTPIM-----DKLSCSNWOKLEAHGLHVG 57

Qy 58 LPEGLMSEVGHNLNIGAGRIYQDIYRINLAIVORNEFTNPQIVASAEAKKSGRLHL 117
Db 64 LPEGLMSEVGHNLNIGAGRIYQDIYRINLAIVORNEFTNPQIVASAEAKKSGRLHL 117

Qy 118 LGLVSDGVSHIDHFLALRAFQKQVFFIHFFADGRDTSPTSGAGLYEQLLQFIAS 177
Db 124 LGLVSDGVSHIDHFLALRAFQKQVFFIHFFADGRDTSPTSGAGLYEQLLQFIAS 177

Qy 178 EKYGELATITGRYYAMDRDKRWRIKMAEYAIYVGGIGOKATVDKAVDVVREYAAQSETDE 237
Db 184 IGGEIATVGRYAMDRDKRWRIKMAEYAIYVGGIGOKATVDKAVDVVREYAAQSETDE 237

Qy 238 FLKPIVF--SDGRYKDDDTLFFNYRADRMQRQICELGLERYKDLNLSVPHKNIQISGM 296
Db 242 FLKPIVGGERRVKDDTLFFNYRADRMQRQICELGLERYKDLNLSVPHKNIQISGM 296

Qy 297 TQYNKEFPFSLPPVTHTNVLAELASQGVTFQHCATEKYPHVTFFNFGGRVQODE 356
Db 301 TQYKTDITFPVAFPPHMGVNVLAELASQGVTFQHCATEKYPHVTFFNFGGRVQODE 356

Qy 357 ERCMVPSPKEVATYDLKPEMNAAGVAEKVEQIESGRHPLVMCNFAPDMVGHGKFEPA 416
Db 361 VRDMIPSPK--VATYDLDPKSAEAVGQKMDRIAGKEFFVNNFAPDMVGHGKFEPA 416

Qy 417 VKAQAQTEAIGKIFACQTYNYVLMYTSDHGNAEKMIAPDGSHTAHTCNLPFTCSSK 476
Db 420 IQGVAATDKAIGYIEACKQGYVLFITADHGNAEMLTEKGTPTKTSHTTNKVPF----- 474

Qy 477 TFEVKSPTPTGDDGKERARALRDVAPTIVQLMGLPVPPEMDGVPLL 522
Db 475 ---IMANAPEGWSLKEGVGLGDVAPTIVLAAMGIEQPEEMSGQNLL 517

RESULT 4
Q8DCW1 PRELIMINARY; PRT; 526 AA.

AC Q8DCW1
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Phosphoglyceromutase.
GN vwl1281.
OS Bacterium vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_taxid=672;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of *Vibrio vulnificus* CMCP6."
RL Submitted (DSC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016801; AA009736.1;
KW Complete proteome.
SQ SEQUENCE 526 AA; 56870 MW; 0ABC0D72A94DF2C0 CRC64;

Query Match 38.6%; Score 1075.5; DB 16; Length 526;
Best Local Similarity 42.9%; Pred. No. 7.3e-80;
Matches 229; Conservative 88; Mismatches 176; Indels 41; Gaps 12;

Qy 6 NVQKVCVLVVIDGWLGSDEQGNATAKAKTPIMDKLSCSNWOKL-EAHGLHVLGPEGLMG 64
Db 18 SAKKPLALVILDGVGREDTASNAIAKATPVMDALIANQNPNTLISAGMDVGLPDGOMG 77

Qy 65 NSEVGHNLNIGAGRIYQDIYRINLAIVORNEFTNPQIVASAEAKKSGRLHLGLVSDG 124
Db 78 NSEVGHNLNIGAGRIYQDIYRINLAIVORNEFTNPQIVASAEAKKSGRLHLGLVSDG 124

Qy 125 GVSHIDHFLALRAFQKQVFFIHFFADGRDTSPTSGAGLYEQLLQFIASEKYGELA 184
Db 138 GVSHIDHFLALRAFQKQVFFIHFFADGRDTSPTSGAGLYEQLLQFIASEKYGELA 184

Qy 185 TITGRYYAMDRDKRWRIKMAEYAIYVGGIGOKA--TVDKAVDVVREYAAQSETDEFLKPI 242
Db 198 SLVGRYYAMDRDNWDRVQVAYDLLTQA--KADFTYDSAVAGLAAAYERGEDEFVKAT 254

Qy 243 VFSDDGR---VKDDDTLFFNYRADRMQRQI-----CECLGLER--YKDLNLSVPHKNI 291
Db 255 EIKAEQGESAMQDGAIVFMNYRADRARQITRTTFVADFAGFERAVFPVNFVM----- 308

Qy 292 QISGMTQYNKEFPFSLPPVTHTNVLAELASQGVTFQHCATEKYPHVTFFNFGGRV 351
Db 309 ---LTQYAADIPAIAPFPASLENTYGEWLSKQGTQLRISETEKTAHVTFNFGGVET 364

Qy 352 QFQDEERCMPSPKEVATYDLKPEMNAAGVAEKVEQIESGRHPLVMCNFAPDMVGHG 411
Db 365 EFAGEERQLVASPK--VATYDLQPEMSSSTELTEKVNAAIKSKYDTIICNYPNADMVGHG 423

Qy 412 KFEPAVKCAQATDEAIGKIFACQTYNYVLMYTSDHGNAEKMIAPD--GSEHTAHTCNLP 470
Db 424 VYEAARAEIAELDASVQGVVEAIEVGGQLLITADHGNAMMVPDPTGGIHTAHTSLPVP 483

Qy 471 FT-CSSTFVFKSPPTGDDGKERARALRDVAPTIVQLMGLPVPPEMDGVPLLE 523
Db 484 LIYVCDKAVFEK-----EGGK-----LSDLAPTNLSLGLAIEPAEMTQGVLVK 526

RESULT 5
Q8FCA6 PRELIMINARY; PRT; 523 AA.

AC Q8FCA6
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE 2,3-bisphosphoglycerate-independent phosphoglycerate mutase (EC 5.4.2.1)
GN Y1BO_OR C4438.
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_taxid=217992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-O6:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*."
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
DR EMBL; AE016768; AAN82874.1;
KW Isomerase; Complete proteome.
SQ SEQUENCE 523 AA; 57234 MW; 7B7A2919BC7FC948 CRC64;

Query Match 38.0%; Score 1059; DB 16; Length 523;
Best Local Similarity 40.9%; Pred. No. 1.6e-78;
Matches 217; Conservative 99; Mismatches 174; Indels 40; Gaps 10;

Qy 8 QOKVCVLVVIDGWLGSDEQGNATAKAKTPIMDKLSCSNWOKL-----LEAHGLHVLGPEGL 62
Db 14 KKPWLVILDGVGREEDQDQNAIFSAKTPVMDAL-----WANRPHTLIDASGLVGLPDQR 69


```
Db 117 IHVGVMSDGVSHIRHKAILEALEG-RNEEFVHCVSDGRDTEPRVFLKYLKVRDF 175
Qy 175 IASEKYGELATIGRYAMDRDRKWERIKMAYEAI-----VGGIGQKATVDKADV---V 226
Db 176 LRTEVGKVASIAGRFYSMDRANNDTELSFRMTTGRREVGG-----DIRSHI 224
Qy 227 RERYAOSSETDEFKLPKIVFSDGRVKDDDTLFFENYRADRMROIQCECLGLERYKDLNSSVP 286
Db 225 CAMYEGLSDTELRLPIDGRIDPKDITIFFNFRADRMROKASFA----ANGNSMIT 280
Qy 287 HPKNIQISGMTQYNKEPPFSLPPTHNLVLAELASQGVTFQFHCATEKYPHVTFEEN 346
Db 281 -----MTEYKKDLGSKVLKICVKNKTLAEVLSRSGIRHSHAENKQAHVTFEEN 331
Qy 347 GGRVOPQDEERCWVSPKAVYDLKPEMNAAGVAKMVEQIESGRHPLVMCNFAPDM 406
Db 332 GGRVOPQDEERCWVSPKAVYDLKPEMNAAGVAKMVEQIESGRHPLVMCNFAPDM 389
Qy 407 VGHGTFEPKAVKCAQATDEAIGKIFEACQTYNTVLMVTSDHGNAEKMIAPDGSSEHTAHC 466
Db 390 VGHGTFEPKAVKCAQATDEAIGKIFEACQTYNTVLMVTSDHGNAEKMIAPDGSSEHTAHC 449
Qy 467 NLVPFTCSSTFFVKTPPTDGDGKERARALRDVAPTVLQMLGVPVPEMDG 518
Db 450 SKVPLIICEBGGV-KASSWGYVDSD--HSLRDVAPTVLEIMGIPRPSMTG 498
```

RESULT 8

```
Q8EW33 ID Q8EW33 PRELIMINARY; PRT; 506 AA.
AC Q8EW33
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Phosphoglycerate mutase.
GN MYPE3740.
OS Mycoplasma penetrans.
OC Bacteria; Firmicutes.
OX NCBI_TaxID=28227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HF-2;
RX MEDLINE=22354719; PubMed=12466555;
RA Sasaki Y., Ishikawa J., Yamashita A., Oshima K., Kenri T., Furuya K.,
RA Yoshino C., Horino A., Shiba T., Sasaki T., Hattori M.;
RT "the complete genomic sequence of Mycoplasma penetrans, an
RT intracellular bacterial pathogen in humans.";
RL Nucleic Acids Res. 30:5293-5300(2002).
DR EMBL; AP004171; BAC44163.1; -.
KW Complete proteome.
SQ SEQUENCE 506 AA; 56436 MW; 915210E83497B289 CRC64;
```

```
Query Match 34.7%; Score 966.5; DB 16; Length 506;
Best Local Similarity 41.2%; Pred. No. 6.5e-71;
Matches 216; Conservative 93; Mismatches 178; Indels 37; Gaps 13;

Qy 11 VCLVVDGWLSDQHGNAIAKAKTPTMDKL-----CSGNWQKLEAHLHVLGPELGM 64
Db 10 VLLAVLDGYSFKDQKGNALNAKTPTMDNLVKEYDHC-----YIEASGEYVGLDQIG 64
Qy 65 NSEVGHNLGAGRIYQDIVRINLAVORNEFVNPQIVASAEAKKSGRHLHLGLVSDG 124
Db 65 NSEVGHNLGAGRIYVGLSLNQDIKTKKFDNSKNTLLEAINAKKNSNIHMLGLSPG 124
Qy 125 GVSHIDHLFALITRAFQKQVPKVIHFADGDRDTSPTSGAGYLEQLLOFIASEKY-GEL 183
Db 125 GVHSEHQHIFEMIRIVSENGL-KPVIHVFGRDVPAPQSISSLERLNDVL--KKYPGTI 181
Qy 184 ATITGRYAMDRKWRERIKMAYEIVGGIGQKATVDKADVVRERYAQSETDFELKPIV 243
Db 182 ATTISGRFYSMDRKRWRERTKQAVDNLGG--ISNNYFNDPIDVYNKQYSENFDFELVPA 239
```

```
Qy 244 FSDGGRV-KDDDTLFFENYRADRMROIQCECL-GLERYKDLNSSVPHF-KNIQISGMTQYN 300
Db 240 INSNVVIKNDADVIAHFDPDRARQISHLFCGSTVYEKND---HPLKNLYIATMITYE 296
Qy 301 KEPPFSLPPTHNLVLAELASQGVTFQFHCATEKYPHVTFEENFGREVGQFQDEBRCM 360
Db 297 GITPTSLTFPTVVVKNFTGFEVANSGLTQLRIATEKYAHVTFEENFGGVEVDLKNESKIL 356
Qy 361 VPSPKVATYDLKPEMNAAGVAKMVEQIESGRHPLVMCNFAPDMVGHGTFEKPEPAYKAC 420
Db 357 VDS-KKVKTYDEVPAMSAVEITDKLIENID--KPDVIVLNFANADMGVGHGTFEKYNEAVLAI 413
Qy 421 QATDEAIGKIFEACQTYNTVLMVTSDHGNAEKMIAPDGSSEHTAHCNLVPTCSSTKTFE 480
Db 414 EALDSQALRIDQKIKELNGTMTFADHGNAEVMDLDDNNPVTKHTTNPVIFISNKKDKVF 473
Qy 481 KSTPPTGDDGKERARALRDVAPTVLQMLGVPVPEMDGVPLEQ 524
Db 474 -----NKPGLSGNVAPTILDFMGLFIPADMDKSLKK 506
```

RESULT 9

```
Q9XE59 ID Q9XE59 PRELIMINARY; PRT; 559 AA.
AC Q9XE59
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Phosphoglycerate mutase (EC 5.4.2.1).
GN PGAM.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Desire;
RA Westram A., Kossmann J.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF047842; AAD24857.1; -.
DR InterPro; IPR006124; Metalloenzyme.
DR Pfam; PF01676; Metalloenzyme; 1.
DR ProDom; PD004429; Pgm_bpd_ind; 1.
DR TIGRFAMs; TIGR01307; pgm_bpd_ind; 1.
KW Isomerase.
SQ SEQUENCE 559 AA; 61257 MW; 68CBEB23585262E0 CRC64;
```

```
Query Match 26.6%; Score 742; DB 10; Length 559;
Best Local Similarity 34.2%; Pred. No. 2.3e-52;
Matches 186; Conservative 92; Mismatches 214; Indels 52; Gaps 16;

Qy 11 VCLVVDGWLSDQHGNAIAKAKTPTMDKLCSG---NWQKLEAHLHVLGIP-BELMGNS 66
Db 22 VAVILVDGWEAKPNEYNAISVAETPVMSLKGAPHEWRHLKAGHNAVGLPTDEDDMGNS 81
Qy 67 EVGHNLGAGRIYQDIVRINLAVORNEFVNPQIVASAEAKKSGRHLHLGLVSDGV 126
Db 82 EVGHNLGAGRIYQDIVRINLAVORNEFVNPQIVASAEAKKSGRHLHLGLVSDGV 139
Qy 127 HSHIDHLFALITRAFQKQVPKVIHFADGDRDTSPTSGAGYLEQLLOFIASEKY---GE 182
Db 140 HSRIDQLLLKGAERGAIRVHALTDGDRVLDGSSVGFMTLENDLAKLRQGGVDAR 199
Qy 183 LATITGR-YVAMDR-DKRWRERIKMAYEIVGGIGQKATVDKADVVRERYAQSETDFELK 240
Db 200 VASGGRMYTMDRYENDWDVVKRGDAQVGLGAPHEKDPVEAIKKLRQEPNDSQYLA 259
Qy 241 PIVESDD-----GRVKDDDTLFFENYRADRMROIQCECLGLERYKDLNSSVPHKNIQIS 295
Db 260 PFVIVDDNGKPVGPILGDGDAVVTNFRADRMVLALEYEDF-DKDFRVRVPK-IHVAG 317
```

[illegible]

[2]
RN SEQUENCE FROM N.A.
RP Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RT "Full-length cDNA from Arabidopsis thaliana."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY085044; AAW61601.1; -;
DR InterPro; IPR006124; Metalloenzyme.
DR InterPro; IPR005995; Pgm_bpd_ind.
DR Pfam; PF01676; Metalloenzyme; 1.
DR ProDom; PD004429; Pgm_bpd_ind; 1.
DR TIGRFAMs; TIGR01307; pgm_bpd_ind; 1.
SQ SEQUENCE 560 AA; 60699 MW; C251A9BFF6AFA07A CRC64;

| | | | | | | | |
|-----|------------|----|--------------|------------|-------------------|-------------------|----------------|
| 13 | LVVIDGWLSD | QY | EHQHNAIAKAKT | PMOKLCSG | ---NMOKLEAHLHVLGP | ---EGLMGNSEV | 68 |
| | 1 | | 1 | 1 | 1 | 1 | 1 |
| | 2 | | 2 | 2 | 2 | 2 | 2 |
| | 3 | | 3 | 3 | 3 | 3 | 3 |
| | 4 | | 4 | 4 | 4 | 4 | 4 |
| | 5 | | 5 | 5 | 5 | 5 | 5 |
| 25 | LIVLDGWESD | Db | PQYNCIHKAPT | PAMDSLKDGP | DTWRLIKAHGTA | VCPLPSEDDMGSEV | 84 |
| | 1 | | 1 | 1 | 1 | 1 | 1 |
| | 2 | | 2 | 2 | 2 | 2 | 2 |
| | 3 | | 3 | 3 | 3 | 3 | 3 |
| | 4 | | 4 | 4 | 4 | 4 | 4 |
| | 5 | | 5 | 5 | 5 | 5 | 5 |
| 69 | GHNLNIGAGR | QY | VIQDVRINLIA | VQRNEFVTPQ | IVASAEAKKGS | GRHLHLGLVSDGGVHS | 128 |
| | 1 | | 1 | 1 | 1 | 1 | 1 |
| | 2 | | 2 | 2 | 2 | 2 | 2 |
| | 3 | | 3 | 3 | 3 | 3 | 3 |
| | 4 | | 4 | 4 | 4 | 4 | 4 |
| | 5 | | 5 | 5 | 5 | 5 | 5 |
| 85 | GHNALGAGRI | Db | YAQCAKLVDL | ALASGKIVDE | GFYISQSEKGT | ---VHLIGLLSDGGVHS | 142 |
| | 1 | | 1 | 1 | 1 | 1 | 1 |
| | 2 | | 2 | 2 | 2 | 2 | 2 |
| | 3 | | 3 | 3 | 3 | 3 | 3 |
| | 4 | | 4 | 4 | 4 | 4 | 4 |
| | 5 | | 5 | 5 | 5 | 5 | 5 |
| 129 | HIDHFLALIR | QY | AFKQLQVPKV | YIHFHFDAG | ROTSPTSGAGY | LBOLOQTASEKY | ---GELA 184 |
| | 1 | | 1 | 1 | 1 | 1 | 1 |
| | 2 | | 2 | 2 | 2 | 2 | 2 |
| | 3 | | 3 | 3 | 3 | 3 | 3 |
| | 4 | | 4 | 4 | 4 | 4 | 4 |
| | 5 | | 5 | 5 | 5 | 5 | 5 |
| 143 | RLDOVOLLL | Db | KGFAGERGA | KRIRVHLLT | DGRVLDGSSV | GFVETLEADLAAL | SKSGVDYAQA 202 |
| | 1 | | 1 | 1 | 1 | 1 | 1 |
| | 2 | | 2 | 2 | 2 | 2 | 2 |
| | 3 | | 3 | 3 | 3 | 3 | 3 |
| | 4 | | 4 | 4 | 4 | 4 | 4 |
| | 5 | | 5 | 5 | 5 | 5 | 5 |


```

QY 127 HSHIDHLFALIRAFKQLQVPKVIHFHFDAGDRPTSPSTGAGYLEQLLOFTASEK-----YGE 182
D 140 HSLRDQVQLLKASERGAIRIRVHILTDGRDVLDCSSGVFVETLESLSQLRDKGIDAR 199
QY 183 LATITGR-YYAMDR-DKRWERIKMAYEA-IVGGIGOKATVDKAVDVVRYRYAQSE-TDEF 238
D 200 IASGGGRMYTMDRYENDMDVVKRGWDAQVILGEAPYK--FQNAVEAVKTLRAETKASDQY 257
QY 239 LKPIVESDD-----GRVKDDDTLIFPNYRADRMQICEGLGLERYKDLN--SSVPHPKNI 291
D 258 LPFVIVDESCKSVGPVVDGDAVVTNFRADRMVLAALAKL---EYADFDFKFDVRYVPK-I 313
QY 292 QISGMYQYNKEFFPFSLF-----PPVTHTNVLAEMLASQGVTPHCA-EETEKYPHVTFFFN 346
D 314 RVAGMLQYDGLKPLSHYLVSPPEIRTS--GEYLVKNGVTRFACRHETVTKFHVTFEWN 371
QY 347 GGREVOFODEERCMPSPKEVA--TYDLKPEMNAAGVAEKMEQIESGRHPLVMCNFAPPD 405
D 372 GNRSGYFDETKKEYEIPSDIGITFNVKPKMKALEIAEKARDAILSGKFDQVRNLPNGD 431
QY 406 MVGHTGKFPAPKACQATDEAIGKIFACQTYNVLMVTSDHGNAEKMIAPDQSEH--- 461
D 432 MVGHTGDIATVYACKAADAENVKILDAIEQVGGIYLVTAHGNADMDVKNKSGOPLD 491
QY 462 -----TAHTCNLVPTCSSKTF-----VFKS-----TPPTGDDGKERARALRDVAPTVLQ 506
D 492 KNGGIQILTSHTLQVPVPAIGGPGHGPVKFRSDIQT-----GLANVAATVMN 540
QY 507 LMGLPVPEMD 517
D 541 FHGFEAPADYE 551

RESULT 14
Q8LD62
ID O8LD62 PRELIMINARY; PRT; 557 AA.
AC O8LD62;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Putative 2,3-bisphosphoglycerate-independent phosphoglycerate
DE mutase.
GN AT1G09780.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RN SEQUENCE FROM N.A.
RA Haas B.J., Volkovsky N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
RL annotation.";
RL Genome Biol. 0:0-0(2002).
RN [2]
RN SEQUENCE FROM N.A.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RT "Full-length cDNA from Arabidopsis thaliana.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Hsuan V.W., Lee J.M.,
RA Quach H.L., Tang C.C., Toriumi M., Wallender E.K., Wong C., Wu H.C.,
RA Yu G., Yuan S., Chen H., Cheuk R., Jones T., Kim C.J., Nguyen M.,
RA Palm C.J., Shinn P., Southwick A., Tripp M.G., Wu T., Davis R.W.,
RA Ecker J.R., Theologis A.;
RT "Arabidopsis Open Reading Frame (ORF) clones.";
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RN SEQUENCE FROM N.A.

```

```

RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Hsuan V.W., Lee J.M.,
RA Quach H.L., Tang C.C., Toriumi M., Wallender E.K., Wong C., Wu H.C.,
RA Yu G., Yuan S., Carninci P., Chen H., Cheuk R., Jones T., Kim C.J.,
RA Ishida J., Jones T., Kamiya A., Kawai J., Kim C.J., Narusaka M.,
RA Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P.,
RA Southwick A., Tripp M.G., Wu T., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RT "Arabidopsis Full Length cDNA Clones.";
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY086182; AAM64261.1; -
DR EMBL: AY150432; AAN12974.1; -
DR EMBL: BT000773; AAN1912.1; -
DR InterPro: IPR006124; Metalloenzyme.
DR InterPro: IPR005995; Pgm_bpd_ind.
DR Pfam: PF01676; Metalloenzyme; 1.
DR ProDom: PD004429; Pgm_bpd_ind; 1.
DR TIGRfam: TIGR01307; Pgm_bpd_ind; 1.
SQ SEQUENCE 557 AA; 60579 MW; D5FC3D06A963B9AB CRC64;

Query Match 25.3%; Score 706; DB 10; Length 557;
Best Local Similarity 34.4%; Pred. No. 2.2e-49;
Matches 189; Conservative 85; Mismatches 213; Indels 62; Gaps 21;

QY 11 VCLVVIDGHLSDSEHGNAIAKAKTPIMDKLCSG---NMQKLEAHGLHVLGP-EGLMGNS 66
D 21 IAVIVLDGGESAPQYNCIHNAPTPAMDSLKHGAPDTWTLKAHGTAVGLPSDEDDMGNS 80
QY 67 EYGHNLNIGAGRVYQDIVRINLAVORNEVTNPQIVASAEARAKKSGRLHLGLVSDGGV 126
D 81 EVGHNLNAGRIFAQAGKALCDQALASGKIFEGEGKPYVSESE--TNTLHLVGLLSDGGV 138
QY 127 HSHIDHLFALIRAFKQLQVPKVIHFHFDAGDRPTSPSTGAGYLEQL-LQFIASEKYG---E 182
D 139 HSLRDQVQLLKASERGAIRIRVHILTDGRDVLDCSSGVFVETLESLSQLRDKGIDAR 198
QY 183 LATITGR-YYAMDR-DKRWERIKMAYEA-IVGGIGOKATVDKAVDVVRE-RYAOSETDEF 238
D 199 IASGGGRMYTLDRYENDMDVVKRGWDAQVILGEAPHK--FKNAVEAVKTLRKEPGANDQY 256
QY 239 LKPIVESDD-----GRVKDDDTLIFPNYRADRMQICEGLGLERYKDLNASSVPHPKNIQI 293
D 257 LPFVIVDESCKSVGPVVDGDAVVTNFRADRMVLAALAKL---EYADFDFKFDVRYVPK-IRY 314
QY 294 SCMTQYNKEFFPFSLF-----PPVTHTNVLAEMLASQGVTPHCAETEKEKYPHVTFFNGSR 349
D 315 AGMLQYDGLKPLSHYLVSPPEIRTS--GEYLVKNGVTRFACRHETVTKFHVTFEWN 372
QY 350 EQVQFOD--EERCMPSPKEVATYDLKPEMNAAGVAEKMEQIESGRHPLVMCNFAPPDVW 407
D 373 SGYFNKELEYEVEIFSDSGI-SFNQVRMKALEIGEKAARDAILSGKFDQVRNIPNGDW 431
QY 408 GHTGFEPAVKACQATDEAIGKIFACQTYNVLMVTSDHGNAEKMIAPDQSEH----- 461
D 432 GHTGDIATVYACKAADAENVKILDAIEQVGGIYLVTAHGNADMDVKNKSGKPALDKE 491
QY 462 -----TAHTCNLVPTCSSKTFV-----FK-----STPPTGDDGKERARALRDVAPTVLQ 508
D 492 GKLIQILTSHTLQVPVPAIGGPGHGPVKFRSDIQT-----GLANVAATVMNLH 540
QY 509 GLPVPPEMD 517
D 541 GFVAPSDYE 549

RESULT 15
Q8H161
ID Q8H161 PRELIMINARY; PRT; 557 AA.
AC Q8H161;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Putative 2,3-bisphosphoglycerate-independent phosphoglycerate
DE mutase.

```


THIS PAGE BLANK (USPTO)